Technical Report 642

Army Maintenance Training and Evaluation Simulation System (AMTESS) Device Evaluation:

Volume I. Overview of the Study Effort

Kenneth W. Unger and Robert W. Swezey
Behavioral Sciences Research Center
Science Applications, Inc.

Robert T. Hays and Angelo Mirabella
Army Research Institute

AD-A159 454

Training and Simulation Technical Area
Training Research Laboratory

U. S. Army
Research Institute for the Behavioral and Social Sciences

July 1984

U. S. ARMY RESEARCH INSTITUTE

FOR THE BEHAVIORAL AND SOCIAL SCIENCES

A Field Operating Agency under the Jurisdiction of the

Deputy Chief of Staff for Personnel

EDGAR M. JOHNSON
Technical Director

L. NEALE COSBY
Colonel, IN
Commander

Research accomplished under contract for the Department of the Army

Science Applications, Inc.

Technical review by

Vanessa Irizarry
Stephen Cormier

Accession For	
NTIS GRA&I	X
DTIC TAB	☐
Unannounced	☐
Justification	
By	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
A-1	

DTIC COPY INSPECTED 3

NOTICES

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER
ARI Technical Report 642

2. GOVT ACCESSION NO.
AD-A159 454

3. RECIPIENT'S CATALOG NUMBER

4. TITLE (and Subtitle)
ARMY MAINTENANCE TRAINING AND EVALUATION SIMULATION SYSTEM (AMTESS) DEVICE EVALUATION: VOLUME I, OVERVIEW OF THE STUDY EFFORT

5. TYPE OF REPORT & PERIOD COVERED
Final Report
January 1982-July 1984

6. PERFORMING ORG. REPORT NUMBER
SAI-84-07-178

7. AUTHOR(s)
Kenneth W. Unger and Robert W. Swezey (BSRC);
Robert T. Hays and Angelo Mirabella (ARI)

8. CONTRACT OR GRANT NUMBER(s)
MDA 903-79-C-0177

9. PERFORMING ORGANIZATION NAME AND ADDRESS
Behavioral Sciences Research Center
Science Applications, Inc.
McLean, VA 22102

10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
2Q263744A795
3.4.4

11. CONTROLLING OFFICE NAME AND ADDRESS
U.S. Army Research Institute for the Behavioral and Social Sciences
5001 Eisenhower Avenue, Alexandria, VA 22333-5600

12. REPORT DATE
July 1984

13. NUMBER OF PAGES
34

14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)
--

15. SECURITY CLASS. (of this report)
Unclassified

15a. DECLASSIFICATION/DOWNGRADING SCHEDULE
--

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

--

18. SUPPLEMENTARY NOTES

This report is Volume I of a three-volume series. Under separate cover are TR 643 (Volume II) and TR 644 (Volume III). TR 645 is available as a related but separate report.

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Simulation
Maintenance training
Transfer of training
AMTESS devices
Training simulators

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This report, Volume I in a three-volume series, provides an overview of a multiyear project whose purpose was to analyze the training effectiveness of two generic maintenance training devices in the Army's AMTESS program. Volume II provides a detailed transfer-of-training evaluation and Volume III presents qualitative information collected during the effort. Results will be used to guide future AMTESS research and development efforts.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

Technical Report 642

Army Maintenance Training and Evaluation Simulation System (AMTESS) Device Evaluation:

Volume I. Overview of the Study Effort

Kenneth W. Unger and Robert W. Swezey
Behavioral Sciences Research Center
Science Applications, Inc.

Robert T. Hays and Angelo Mirabella
Army Research Institute

Submitted by
Robert J. Seidel, Chief
Training and Simulation Technical Area

Approved as technically adequate
and submitted for publication by
Harold F. O'Neil, Jr., Director
Training Research Laboratory

U.S. ARMY RESEARCH INSTITUTE FOR THE BEHAVIORAL AND SOCIAL SCIENCES
5001 Eisenhower Avenue, Alexandria, Virginia 22333-5600

Office, Deputy Chief of Staff for Personnel
Department of the Army

July 1984

Army Project Number
2Q263744A795

Training and Simulation

ARI Research Reports and Technical Reports are intended for sponsors of R&D tasks and for other research and military agencies. Any findings ready for implementation at the time of publication are presented in the last part of the Brief. Upon completion of a major phase of the task, formal recommendations for official action normally are conveyed to appropriate military agencies by briefing or Disposition Form.

FOREWORD

The Army Maintenance Training and Evaluation Simulation System (AMTESS) program, administered by the Army's Project Manager-Training Devices (PM-TRADE), is intended to develop a family of devices which can be used to train personnel in tasks required by a range of Military Occupational Specialties (MOS). The Army Research Institute (ARI) is evaluating this program in support of PM-TRADE.

Previous ARI reports have examined the features required by such a training simulator and the type of analysis needed to set up a testing program for the device. This report gives the results of quantitative and qualitative field testing of two prototypes. These results should prove valuable in designing future maintenance simulators.

The next step in the AMTESS program will be laboratory research at George Mason University in Fairfax, Virginia. This research, performed under ARI contract, will test the AMTESS devices under scientifically controlled circumstances.

Edgar M. Johnson

EDGAR M. JOHNSON
Technical Director

ACKNOWLEDGMENTS

The authors wish to acknowledge the special contributions made to this project by the following persons: Martha J. Cole, who served as the project data collector at Fort Bliss, Texas, and who helped to draft preliminary versions of this document; M. Joseph Moyer, who served as the data collector at Aberdeen Proving Ground, Maryland; Peter Baker of the U.S. Army Air Defense School, who served as the project liaison individual at Fort Bliss; James Dees, who provided similar services at Aberdeen Proving Ground; Mr. Tom Houston, and Dr. Garvin Holman of ARI, who served as project monitors for this effort during its early phases; and Mr. Michael Singer of ARI, who served on the technical team.

ARMY MAINTENANCE TRAINING AND EVALUATION SIMULATION SYSTEM (AMTESS) DEVICE EVALUATION: VOLUME I, OVERVIEW OF THE STUDY EFFORT

EXECUTIVE SUMMARY

Requirement:

The objective of the AMTESS program is to provide cost- and training-effective maintenance simulators that can be easily adapted to a variety of maintenance tasks across a number of MOS. During Phase I of this program, four conceptual versions of AMTESS devices were developed. During Phase II, two prototype AMTESS devices were fabricated. The objective of this report is to summarize quantitative (transfer-of-training) and qualitative (user opinion) data about the two prototype devices. Volumes II and III of this series provide detailed information on quantitative and qualitative results, respectively.

Procedure:

Quantitative data were obtained by training students at Fort Bliss, Texas, and Aberdeen Proving Ground, Maryland, to perform electronic or mechanical maintenance tasks using either conventional instructional methods (lectures and "hands-on" experience) or one of the two AMTESS simulators. These students were subsequently tested on their ability to perform these tasks on operational equipment. Qualitative data were obtained by administering questionnaires to students, instructors, and course developers; and by analyzing observations made by the on-site data collector during the course of the transfer-of-training study. Quantitative and qualitative data were analyzed and combined in order to draw conclusions about the devices.

Findings:

Analysis of quantitative data revealed that both AMTESS devices provided an acceptable level of training in each of two widely different maintenance contexts, encompassing both mechanical and electronic maintenance training situations. In all cases, students trained on the prototype simulators were able to pass the Army School's criterion for the chosen tasks. Results showed that the devices were 87% to 98% as effective as conventional training where "percent steps passed" was the test measure. However, for "time to complete task" and "data collector interventions," the devices were considerably less effective than conventional training. Conventionally trained students needed only 52% to 86% as much time to finish the tasks and a mean of 78% as many data collector interventions to finish the tasks. In other words, the "device" students learned how to perform their tasks but were notably slower and required more supervision than conventionally trained students. Device training was least effective for students who were the least experienced.

Analysis of qualitative data revealed that the utility of specific device features varied with the tasks that the devices trained. In general, students, instructors, and course developers praised a variety of device features including feedback provided to students, the ease with which malfunctions can be inserted into the devices, and the quality of audio and visual stimuli presented to students. The low reliability and durability of the devices detracted from their overall effectiveness.

Utilization of Findings:

Simulators developed from the prototypes could be effective, as part-task trainers, to support conventional training. As designed, however, they should not substitute for such training, except where equipment is dangerous, subject to severe damage, or so inaccessible that routine instruction is impractical. Conventional training may be especially necessary for inexperienced soldiers and for tasks requiring removal and replacement of parts. More definitive conclusions about the devices are difficult to make because only gross comparisons between device and conventional training were possible. However, planned follow-up research (with better experimental controls than were possible at the field locations) will lead to specific conclusions about how to design and use generic simulation systems for different types of maintenance tasks, types of trainee, and training strategies.

ARMY MAINTENANCE TRAINING AND EVALUATION SIMULATION SYSTEM (AMTESS) DEVICE EVALUATION: VOLUME I, OVERVIEW OF THE STUDY EFFORT

CONTENTS

LIST OF TABLES

LIST OF FIGURES

ARMY MAINTENANCE TRAINING AND EVALUATION SIMULATION SYSTEM (AMTESS) DEVICE EVALUATION: VOLUME I, OVERVIEW OF THE STUDY EFFORT

INTRODUCTION

The purpose of this report is to summarize quantitative (transfer-of-training) and qualitative (user opinion) findings about two prototype Army Maintenance Training and Evaluation Simulation System (AMTESS) devices. The objective of the AMTESS program is to provide cost- and training-effective maintenance simulators designed to be easily adapted to a variety of MOSs and maintenance tasks (Dybas, 1981, 1983). The AMTESS program has evolved through several phases of development. During Phase I, plans for four conceptual versions of an AMTESS device were developed. Two of these plans were selected for further development during Phase II. Consequently, two prototype AMTESS devices were constructed to demonstrate the alternative approaches. These "breadboard" (or prototype) devices are working models capable of displaying the features and concepts proposed by the development contractors. The two prototype devices were evaluated in a field transfer-of-training study.

This overview document (Volume I) is supplemented by two detailed technical volumes. Volume II addresses the field transfer-of-training study of the two prototype AMTESS devices, and Volume III presents detailed qualitative data collected during the course of the effort. Other relevant reports in the series include: Criswell, Unger, Swezey, and Hays (1983) which addresses the history of the AMTESS program and the features of the AMTESS devices, Woelfel, Duffy, Unger, Swezey, Hays, and Mirabella (1984) which deals with AMTESS front-end analysis activity, and Criswell, Swezey, Allen, and Hays (1984) which deals with human factors aspects of the devices.

The evaluation of the two prototype AMTESS devices is not typical of training device evaluation studies. These devices are products of an ongoing research and development effort. They were designed to demonstrate conceptual approaches to hardware, software, and courseware design. The purpose of the evaluation was to determine if these conceptual approaches warranted further development. (No attempt was made to determine directly which of the two devices was superior to the other.) In order to determine if the devices merited further development, the evaluation was designed to provide data on the following questions:

- Do the curricula associated with the devices provide adequate training.
- How does the training effectiveness of the AMTESS devices compare to the training effectiveness of conventional methods?
- Is the AMTESS concept of modular training simulators feasible?

In order to answer these questions, both quantitative and qualitative data were collected. Quantitative data were collected to compare the training effectiveness of the devices to the training effectiveness of the conventional

methods in order to assess the overall level of training provided by the devices. Qualitative data were collected from knowledgeable individuals to assess the feasibility of device concepts, and the utility of specific device features.

A transfer-of-training paradigm was used to assess the training effectiveness of each device. This type of study involved the use of two groups of students: an experimental group that received simulator training prior to performance testing on operational equipment, and a conventionally trained group that received conventional Army instruction in the appropriate MOSs including a "hands-on" component using actual operational equipment prior to performance testing on operational equipment. If the two groups are equated on other relevant factors, their differences in performance on operational equipment can be attributed to the influence of training received by the two groups.

The evaluation of the two prototype devices involved two sites and students from eight MOSs. In order to improve the clarity of communication, the evaluation effort is presented as six separate experiments. Figure 1 presents the overall design of these six experiments. Brief descriptions of the six experiments are presented in a subsequent section of this volume of the report. Detailed descriptions of the experiments are presented in Volumes II and III of the report.

DESCRIPTION OF SIMULATORS

The prototype AMTESS devices were developed independently by two contractors: Grumman Aerospace Corporation, and a consortium of Seville Research Corporation and Burtek, Inc. Each prototype maintenance trainer was designed to train two sets of maintenance tasks: one involving mechanical maintenance in the automotive area and the other involving electronic maintenance for a portion of a radar system.

Both breadboard devices incorporate a microcomputer-based system attached to two-dimensional (2-D) cathode ray tubes (CRTs), and to uniquely designed three-dimensional (3-D) modules simulating portions of the automotive and radar system hardware used for "hands-on" training. The core unit (i.e., the computer system and CRTs) were designed to remain constant across various types of maintenance training activity. The 3-D module (and the accompanying lesson materials and software) were designed to be removed and replaced as required for training on various maintenance tasks and MOSs.

Beyond these overall AMTESS requirements, the two contractors used different approaches in implementing their concepts. The Seville/Burtek device was designed to be used primarily to replace operational equipment in training, whereas the Grumman device was designed to supplement training on operational equipment. A second major difference concerns the way information is presented on the two devices. The Seville/Burtek device utilizes a 35mm rear projection visual display system, whereas the Grumman device is videodisc-based. More complete descriptions of these devices are presented below.

EXPERIMENT	MOS	TRAINING	TESTING
1	63B30	Seville/Burtek Simulator N=20	Operational Equipment
	63B30	Conventional N=20	
2	63W10	Seville/Burtek Simulator N=21	Operational Equipment
	63W10	Conventional N=20	
3	63H30	Grumman Simulator N=10	Operational Equipment
	63H30	Conventional N=12	
4	63D30	Grumman Simulator N=12	Operational Equipment
	63D30	Conventional N=11	
5	24C10	Seville/Burtek Simulator N=10	Operational Equipment
	24C10	Conventional N=12	
6	24E 24G 24R	Grumman Simulator N=10	Operational Equipment & Simulator

Figure 1. Overall design of the transfer-of-training experiments.

Seville/Burtek

The Seville/Burtek AMTESS device consists of four workstations:

1. an instructor station

2. a student station

3. a 3-D module that emulates a Cummins NHC-250 diesel engine

4. a 3-D module that emulates an Improved Hawk High Power Illuminator Radar transmitter

These components are connected by interface cables.

The instructor station consists of an instructor panel for input of frequently used commands, a CRT for displaying the status of the 3-D module and editing lessons, a keyboard for input of infrequently used commands, a printer, and a desk housing the computer, input/output equipment, floppy and hard disk drives.

The instructor controls the training process through the instructor panel. He or she selects an exercise by pushing a few buttons on this panel. Once an exercise has been selected, the 3-D module will display the symptoms of the selected malfunction. The instructor may examine student performance directly on the instructor CRT or on a printout either as the student works on a lesson or after the student has completed the lesson. The student performance record includes the exercise name and number, the action the student took (time the action was taken), exercise program step number, and information concerning the appropriateness of student actions (deviation, bad order, etc.). The instructor may use the CRT and the keyboard to modify messages and slides presented to the student, solutions to exercises, and other pertinent aspects of lessons.

The student station consists of a responder panel, a CRT, and a slide projector unit. The student CRT presents questions, provides feedback, and gives direction to students. Students use the responder panel to indicate answers to questions (yes, no), enter corrective procedures (service, remove/replace, inspect), indicate faulty components (starter motor, alternator, batteries), and request instructor assistance. The slide projector unit presents diagrams, flow charts, and other pertinent information.

One of the 3-D modules of the Seville/Burtek device is a full-scale reproduction of a Cummins NHC-250 diesel engine. This 3-D module is composed of engine components (i.e., starter motor, batteries, etc.), controls, displays, and test equipment.

A second 3-D module of the Seville/Burtek device is a full-scale reproduction of an Improved Hawk High Power Illuminator Radar transmitter unit. This module is composed of a cabinet, interior components (i.e., cables, power supplies, etc.), controls, displays, and test equipment.

Table 1 summarizes the features of the Seville/Burtek device. Garlichs, Miller, and Davis (1983) have provided a detailed discussion of the design and features of the Seville/Burtek device.

Table 1

Description of Seville/Burtek Device Features

Feature	Description
Remove/replace capability	Components of the 3-D module can be removed and replaced by students.
Student CRT unit	Text presented on this CRT directs student actions and provides feedback.
Student responder unit	Students turn a thumbwheel and push buttons to enter their decisions (i.e., inspect hoses) into the training device.
Instructor CRT unit	Instructors are able to monitor student actions on this CRT unit.
Hardcopy printout	A record of student performance can be printed as the student performs the lesson or after the lesson is completed.
Lesson arrangement	The sequence in which students participate in lessons can be controlled by the instructor.
Editing system	Text, graphics, and all procedures involved in a lesson can be modified by instructors.
Malfunction insertion	Numerous faults can be inserted into the simulator by manipulating two controls on the instructor station.
Random malfunction insertion	This feature randomly selects a malfunction and inserts it into the simulator.
Slide projector unit	Photographs, diagrams, and other visual aids are presented via slide projector.
Sound effects	Various equipment sounds are simulated.

Grumman

The Grumman AMTESS device consists of four workstations:

1. an instructor station
2. a student station
3. a 3-D module that emulates the starting and charging systems of an M110A2 self-propelled howitzer
4. a 3-D module that emulates an Improved Hawk High Power Illuminator Radar transmitter

These components are connected by interface cables.

The instructor station consists of a CRT, keyboard, and printer. These components are located on a desk which houses a computer, floppy and hard disk drives, and videodisc system. Included in the configuration is the capability to display in hard copy a student performance record for each lesson segment which includes lesson segment number, total number of errors committed, total time required to complete a segment, and a performance index based on total time required to complete a segment. The instructor enrolls students, assigns lesson segments, edits, or creates new lesson segments, and requests hard copy printouts of student records after a segment has been completed. This is accomplished via a keyboard and instructor CRT. When students are participating in lesson segments, the instructor CRT displays event numbers which correspond to actions that a student has taken during the course of a segment.

The student station consists of a color CRT with an associated touch panel. This station queries students, provides directions and feedback to students. Information is presented in the form of still and moving video frames, computer-generated text, and an audio track. Students use the touch panel to answer questions, enter corrective procedures, identify faulty components, and to request help.

One of the 3-D modules of the Grumman device consists of controls, displays, test equipment, and components which simulate the starting and charging system for the M110A2 self-propelled howitzer.

A second 3-D module of the Grumman device is a full-scale reproduction of an Improved Hawk High Power Illuminator Radar transmitter unit. The 3-D module is composed of a cabinet, interior components (cables, power supplies, etc.), controls, displays, and test equipment.

Table 2 summarizes the features of the Grumman device. Campbell, Stonge, Cothran, Anaya, and Sicuranza (1980) have provided a detailed description of the Grumman device design and features.

Table 2

Description of Grumman Device Features

Feature	Description
Video disc system	Still frames, motion frames, and computer-generated text are presented by this system on the student CRT unit in order to direct student actions and provide feedback.
Touch panel	Students enter their decisions into the training device by touching certain locations on the student CRT that display words (yes, no, etc.), pictures (master switch, instrument switch, etc.) or schematics.
Request help	Many frames presented by the videodisc player allow students to request help by touching the CRT. Help is presented in the form of audio and/or visual cues.
Repeat lesson	At certain predetermined points in the training program, students may choose to repeat segments, or parts of segments.
Call instructor	When students make two consecutive errors, the device ceases to accept student input and the student receives a message to call an instructor.
Sound effects	Various engine sounds (cranking, idle, shut-down) are simulated in the M110A2 howitzer configuration.
Hardcopy printout	A record of student performance can be printed after a lesson is completed.
Automated pre-lesson check	Prior to starting a lesson, the device checks to ensure that all switches, cables, etc. of the 3-D module are in the correct configuration. Instructions for correcting erroneous configurations are presented on the student CRT.
Lesson arrangement	Students normally complete lessons in a fixed sequence. The training device keeps track of the segments that a student has completed, presenting the appropriate segment each time a student works with the device.
Universal instructor	This feature allows instructors to present segments in any sequence; however, no record of student performance is kept when this feature is enabled.
Instructor CRT	Information about the videodisc system is presented on the instructor CRT when students participate in lessons on the simulator.

METHOD

Evaluation of the AMTESS devices took place at Aberdeen Proving Ground, Maryland (APG) for the automotive configuration and at Fort Bliss, Texas for the electronics configuration. Seville/Burtek's 3-D automotive module simulates a Cummins NHC-250 diesel engine, whereas Grumman's 3-D automotive module simulates the starting and charging system of an M110 self-propelled howitzer. Since two different vehicles are simulated by the devices, it was necessary to include students from two MOSs at APG in the evaluation study. (No single MOS involves maintenance activity for both types of vehicles.) Thus, students from MOS 63W10 (Direct Support Vehicle Repairman) were included in the evaluation of the Seville/Burtek simulator, and students from MOS 63D30 (Self-propelled Field Artillery Systems Mechanic) were included in the evaluation of the Grumman simulator.

Since both the Seville/Burtek and Grumman simulators addressed the Improved High Power Illuminator Radar transmitter of the Hawk missile system, students in training for MOS 24C10 (Hawk Missile Firing Section Mechanic) at Fort Bliss were used in the evaluation of both devices.

Thus, three experiments were planned for the evaluation - one per MOS. This plan was modified when it appeared that an insufficient number of students was available to participate in the evaluation, and when other constraints were encountered.

Evaluation of the Seville/Burtek device at APG was expanded from one to two experiments with the inclusion of 63B30 (Organizational Maintenance Supervisor) students. The 63B30 students were required because, upon receipt of the simulator, it was discovered that most of the tasks included in the simulator curriculum addressed tasks performed by 63B30 students.

Evaluation of the Grumman device at APG was expanded from one to two experiments with the inclusion of 63H30 (Direct Support Maintenance Supervisor) students. These students were included in the evaluation because the availability of students in MOS 63D30, for which the device was originally designed, was low.

Thus, four separate experiments were conducted at APG: Experiments 1 and 2 addressed the Seville/Burtek device while Experiments 3 and 4 addressed the Grumman device.

Two experiments were conducted at Fort Bliss. Experiment 5 addressed the evaluation of the Seville/Burtek device and involved MOS 24C10 (Improved Hawk Master Mechanic) students. A single group approach was used to evaluate the Grumman device at Fort Bliss. It was not possible in this situation to make direct comparisons between training conditions because the Air Defense School was unable to provide students for a conventional training condition. The study involved students in MOSs 24E (Improved Hawk Fire Control Mechanic), 24G (Improved Hawk Information and Coordination Control Mechanic), and 24R (Improved Hawk Master Mechanic) and is referred to herein as Experiment 6.

Both Army Schools had previously developed relevant job performance tests for assessing training mastery. Students must be able to pass the

tests in order to demonstrate satisfactory maintenance performance for course graduation. These tests, which were comprised of GO/NO-GO checklists (i.e., either a student had performed each step or he had not performed it), were to be used as the primary transfer-of-training criterion measures. These tests were, however, found to be inadequate because they were not sufficiently detailed for the purpose of a transfer-of-training evaluation. Therefore, new and expanded versions of these tests were developed.

The three primary criterion measures employed in the transfer-of-training studies were the percentage of performance test steps passed by each student, the time required to complete each task and the number of interventions which were required on the part of the data collector to ensure that safe working conditions were maintained during the testing. Experimental to conventional (E/C) and conventional to experimental (C/E) ratio scores and scores for reclustered tasks (defined later in the report) were also computed. The distribution of these criterion measures across the six experiments is presented in Table 3. Volume II of this series provides detailed information concerning the six transfer-of-training experiments.

Six different categories of opinion-based data were also collected during the course of transfer-of-training experiments. These data, discussed in detail in Volume III of the series, included:

1. Initial instructor questionnaires, which were filled out after instructors were trained to use the simulators but before they had used the devices to train students.

2. Instructor questionnaires, which were filled out after instructors had used the devices to train students.

3. Course developer questionnaires, which were completed after the respondents had extensive experience operating the devices.

4. Trainee questionnaires, which were completed by students after they completed training with the devices and were tested on operational equipment.

5. Structured interviews, which were conducted with various individuals who were knowledgeable about the simulators.

6. Observations, which were made by the on-site data collectors about specific device features and lessons during the course of the transfer-of-training studies. Data collectors also assessed the reliability of the devices and other aspects of trainee-simulator interactions.

The qualitative data that were recorded during each experiment are presented in Table 4.

Brief descriptions of each of the six experiments are presented below.

Table 3

Criterion Measures Employed in AMTESS Field Evaluation Studies

	EXPERIMENT					
	1	2	3	4	5	6
Device	S/B[a]	S/B	G[b]	G	S/B	G
MOS	63B30	63W10	63H30	63D30	24C10	24E, G, R
Location	APG[c]	APG	APG	APG	Bliss[d]	Bliss
Sample Size						
Simulator Group	20	21	10	12	10	10
Conventional Group	20	20	12	11	12	0
Number of Tasks Tested	4	5 subtasks	8 subtasks	8 subtasks	4	8
Criterion Measures						
E/C and C/E Ratios	X	X	X	X	X	
% Steps Passed	X	X	X	X	X	X
Time to Complete	X	X	X	X	X	X
Data Collector Interventions	X	X	X	X		
Other					Instructor Ratings School Exams	Case Study Approach
Number of Reclustered Tasks Tested	6	5	6	6		
Criterion Measures						
% Steps Passed	X	X	X	X		
Data Collector Interventions	X	X	X	X		

[a]S/B = Seville/Burtek

[b]G = Grumman

[c]APG = Aberdeen Proving Ground, MD

[d]Bliss = Fort Bliss, TX

Table 4

AMTESS Field Evaluation Studies Qualitative Data

	Experiment					
	1	2	3	4	5	6
DEVICE	S/B[a]	S/B	G[b]	G	S/B	G
MOS	63B30	63W10	63H30	63D30	24C10	24E, G, R
LOCATION	APG	APG	APG	APG	BLISS	BLISS
SAMPLE SIZE:						
Initial Instructor Questionnaire	2	3	5	-	11	2
Instructor Questionnaire	1	1	6	-	1	-
Course Developer Questionnaire	2	-	3	-	-	1
Trainee Questionnaire	20	67	15	15	10	10
Structured Interviews	4	1	5	-	2	5
OTHER DATA	DCO[c]	-	DCO	-	DCO	DCO

[a]Seville/Burtek

[b]Grumman

[c]Data Collector Observations

Experiment 1

The first experiment involved the Seville/Burtek device at APG. Twenty (20) MOS 63B30 students were trained to perform four maintenance tasks with the simulator, while an additional 20 were trained with conventional (lecture and "hands-on" experience) methods. These tasks included troubleshooting an oil pump failure (organizational level only), adjusting alternator drive belts, removing/replacing a starter motor, and inspecting an electrical system. Training for both groups of students took place in groups of two.

All students were tested on their ability to perform the four maintenance tasks on an M809 series 5-ton truck. Testing was conducted individually for all students. All testing was conducted by the on-site data collector in a single session (when possible) within 24 hours after training had been completed.

Prior to filling out questionnaires or participating in structured interviews, respondents were briefed on the purpose of the questionnaires (or interviews) and the types of questions that would be asked. Initial instructor questionnaires were completed in a group setting while all other questionnaires were completed independently.

Experiment 2

The second experiment also involved the Seville/Burtek device at APG. In this experiment, 20 students from MOS 63W10 were trained to troubleshoot, remove and replace an oil pump using conventional methods, while 21 students were trained to perform this task using the simulator. The task is composed of the following five subtasks: perform organizational troubleshooting, perform direct support troubleshooting, remove the oil pump filter, remove the oil pump, and replace the oil pump and filter. All students were tested on their ability to perform these subtasks on an M809 5-ton truck. Procedures for collecting opinion-based data were similar to those reported in Experiment 1. Training and testing procedures were similar to those in Experiment 1.

Experiment 3

This experiment involved an assessment of the Grumman device at APG. MOS 63H30 students were trained to perform eight subtasks involving a defective voltage regulator on an M110A2 self-propelled howitzer. Ten students were trained with the simulator, while 12 students were trained conventionally. The subtasks were as follows: confirm malfunction, troubleshoot electrical system, perform vehicle test meter (VTM) hook-up and check-out, perform generator regulator charging circuit test, troubleshoot charging circuit, remove/replace generator voltage regulator, perform VTM hook-up and check-out, perform generator regulator charging circuit test.

All students were tested on their ability to perform these subtasks on the howitzer. Test procedures employed during the evaluation of the Grumman

device were similar to those used in the evaluation of the Seville/Burtek device, i.e., all students were tested individually; all testing was conducted by the on-site data collector; testing was conducted within 24 hours after training had been completed; and testing was conducted in a single session (where possible). Procedures for collecting opinion-based data were similar to those for the previous experiments.

Experiment 4

Experiment 4 also involved an evaluation of the Grumman device at APG. In this experiment, 12 MOS 63D30 students were trained to perform eight subtasks (identical to those reported in Experiment 3) using the simulator, while 11 students were trained using conventional methods. Procedures followed in this experiment were identical to those followed in Experiment 3.

Experiment 5

This experiment involved an evaluation of the Seville/Burtek simulator at Fort Bliss, Texas. Students from MOS 24C10 were trained on several Hawk system maintenance tasks, using simulator-based or conventional training methods. Students were trained in groups of two. The students were then tested on their ability to perform a subset of these tasks on operational Hawk radar equipment. Testing was conducted on an individual basis by School personnel.

Constraints imposed by the U.S. Army Air Defense School seriously degraded the extent to which a controlled experimental design could be implemented at Fort Bliss. Due to School requirements, the number and types of problems upon which data could be collected could not be placed under experimental control. Further, students in the conventionally trained group were trained in a so-called "lockstep" fashion (where the primary mode of instruction is lecture-based, and where entire classes move through the curriculum en masse, as opposed to individual student pacing); whereas simulator-trained students were trained in a self-paced format. The results of transfer-of-training data reported in Experiment 5 are limited by this major confound as well as by a variety of additional constraints encountered at Fort Bliss.

Opinion-based data collection forms and procedures used during Experiment 5 were highly similar to those used in the other experiments.

Experiment 6

A single group case study approach was used to evaluate training provided by the Grumman device at Fort Bliss, Texas. It was not possible to make direct comparisons between simulator-trained students and conventionally trained students in Experiment 6 because the Missile School at Fort Bliss was unable to provide students for a conventional training condition. Thus, the data reported in this experiment are primarily descriptive.

Opinion-based data collection forms and procedures used during Experiment 6 were also similar to those used in the other experiments.

RESULTS

As previously noted, both quantitative and qualitative data were collected during the course of the evaluation. Quantitative data address the performance of students who were trained with the simulator or conventional methods, and who were subsequently tested on their ability to perform troubleshooting tasks on operational equipment. These data are discussed in detail in Volume II of this series of reports. Qualitative data address the opinions of instructors, students, course developers, and other knowledgeable individuals towards the AMTESS devices in general, and towards specific features of those devices. These data are discussed in detail in Volume III of this series of reports.

Quantitative Data

An analysis of student background data was conducted to determine if differences existed between the simulator-trained and the conventionally trained students prior to the beginning of training. In terms of age, grade, and Armed Services Vocational Aptitude Battery (ASVAB) (i.e., general maintenance, mechanical maintenance, general technical, and electronics) scores, statistical tests revealed no significant differences between the students assigned to the conventional training groups and those assigned to the simulator-based training groups.

The percentage of performance test steps passed, the length of time required to complete each test, and the number of data collector interventions required to ensure safe completion of each task were the three dependent variables of interest in the between-group training effectiveness comparisons. Prior to testing the statistical significance of differences between the two training conditions, an effort was made to assess the practical importance of the magnitude of the differences between the two conditions. For each of the three variables described above, the performance of the two groups of students was compared by dividing the scores of one group by the scores of the other group and multiplying the result by 100. The value which results, known as E/C (or C/E) ratio indicates the extent to which the performance of one group (experimental or conventional) approaches that of the other. Percentage of performance tests steps passed scores were compared by computing E/C ratios. Experimental/conventional ratios greater than 100 indicate that performance of the experimental (i.e., simulator-trained) group exceeds that of the conventionally trained group. Since low scores indicate superior performance for the time to complete task measure and the data collector intervention measure, conventional (C) scores were divided by experimental (E) scores, and the result was multiplied by 100. The resulting value indicates the extent to which conventional group performance approaches experimental group performance. (A C/E score of 95 for data collector interventions indicates that conventionally trained students required 95% of the number of data collector interventions required by experimentally trained students.) Conventional/experimental ratio values exceeding 100 indicate superior performance by the experimentally trained students.

E/C and C/E ratios were computed for Experiments 1 through 5. (These ratios could not be computed for Experiment 6 because data were not available for conventionally trained students.) Mean E/C and C/E ratios are presented in Table 5.

Inspection of Table 5 indicates very high E/C ratios for the percentage of performance test steps passed measure for all five experiments. These figures indicate that students who were trained on the AMTESS devices passed almost as many steps on the performance test as students who were trained conventionally. Lower C/E ratios were obtained for the time to complete task measure and the data collector intervention measure.

Table 5

Mean E/C and C/E Ratios for Data Collected during AMTESS Evaluation

Experiment Device MOS	1 S/B[a] 63B30	2 S/B 63W10	3 G[b] 63H30	4 G 63D30	5 S/B 24C10	Grand Mean
E/C ratio for percent steps passed	98	87	95	91	89	92
C/E ratio for time to complete task	72	52	86	74	75	71.8
C/E ratio for data collector interventions	71	49	131	60		77.8
E/C ratios for instructor ratings					85	
E/C ratios for school administered exams					99	

[a]S/B = Seville/Burtek

[b]G = Grumman

- The lowest E/C and C/E ratios were obtained by the 63W10 students (Experiment 2). The low scores obtained by these highly inexperienced students may be due to their lack of familiarity with Army Technical Manuals (TMs), tools, equipment, etc. If this is the case, then it may be appropriate to use the AMTESS device as an adjunct to conventional training rather than as a replacement for conventional training for inexperienced students.

For each of the first five experiments, a series of t-tests (or Mann-Whitney U-tests depending on sample size) was conducted to test for differences between conventionally trained and simulator-trained groups in terms of percentage of performance test steps passed, time to complete task, and number of data collector interventions. Since there was no reason to expect that one group would perform better (or worse) than the other group, two-tailed tests were conducted.[1]

To determine if differences in specific skills and knowledges existed between the two groups, the tasks performed by students were reclustered into a set of more homogeneous tasks, i.e., TM selection, mechanical inspection, remove/replace, etc.[2] Data based upon these reclustered tasks were analyzed in terms of percentage of performance test steps passed and number of data collector interventions.

Table 6 presents the percentage of significance tests which indicated superior performance by the conventionally trained students across five experiments.

Qualitative Data

Responses to initial instructor questionnaires, instructor questionnaires, course developer questionnaires, trainee questionnaires, structured interviews and data collector observations are summarized in Table 7. In general, respondents in the radar and automotive MOSs made similar comments about the devices. There are notable exceptions however:

1. For both devices, respondents at Fort Bliss emphasized the simulators' capability to provide hands-on practice to a greater extent than did respondents at APG.

2. For the Grumman device, respondents at APG commented on the value of the videodisc system to a greater extent than did respondents at Fort Bliss.

---

[1]Rationale for the statistical tests employed in this study is discussed in Volume II of this series.

[2]This reclustering was originally developed by Mirabella and Holman and reported in Evans and Mirabella (1982).

3. The 63B30 students at APG reported confusion with respect to the number of stimuli to which they were required to attend. This negative comment was not made by 63W10 students or by students at Fort Bliss.

Table 6

Percentage of Significance Tests Indicating Significantly Superior Performance by Conventionally Trained Students across Five Experiments

Experiment Device MOS	1 S/B[a] 63B30	2 S/B 63W10	3 G[b] 63H30	4 G 63D30	5 S/B 24C10	Mean
Percentage of performance test steps passed	0	40	13	38	25	23
Time to complete task	75	100	13	25	50	53
Number of data collector interventions	25	20	13	12	-	18
Percentage of performance test steps passed for reclustered tasks	17	60	33	25	-	34
Number of data collector interventions for reclustered tasks	33	60	17	17	-	32
Instructor ratings	-	-	-	-	20	-
School administered exams	-	-	-	-	0	-
Mean	30	56	18	24	24	

[a]S/B = Seville/Burtek

[b]G = Grumman

The information presented in Table 7 does not provide an assessment of the conceptual approaches undertaken by the two device manufacturers. The purpose of the following section of the report is to integrate the qualitative data from all six experiments in order to consider benefits and liabilities of the conceptual approaches utilized by Grumman and by Seville/Burtek in their respective devices.

Table 7

Summary of Positive and Negative Statements about the AMTESS Devices

Experiment 1 (MOS 63B30) Seville/Burtek Device	
Positive:	Respondents hold favorable opinions of the simulator.
	Ease of inserting malfunctions is valuable.
	Performance monitoring is valuable.
	The simulator is safer than operational equipment.
Negative:	Students were confused by the materials to which they must attend
	Physical fidelity of the 3-D module is too low for certain remove/ replace tasks.
	The reliability and durability of the device should be increased.
	Lessons did not include STE/ICE (test equipment) set-up and check-out.
	The 3-D module records normal vibrations as student errors.
Experiment 2 (MOS 63W10) Seville/Burtek Device	
Positive:	Respondents hold favorable opinions of the simulator.
	Respondents liked features including feedback, proceduralized instructions, slide projector unit, malfunction insertion.
Negative:	Lessons did not include STE/ICE set-up and check-out.
	The 3-D module records normal vibrations as student errors.
	The durability of the device is low.
Experiment 3 (MOS 63H30) Grumman Device	
Positive:	The ability to perform troubleshooting tasks on the 3-D module is a valuable feature.
	The videodisc system is an effective motivating feature.
Negative:	The device frequently malfunctions.
	Lessons are inflexible.
	Some lessons are too simple or inappropriate.
	The student performance record is of little value.
	System response time is too slow.
Experiment 4 (MOS 63D30) Grumman Device	
Positive:	Students liked features including the 3-D module, the request help feature, the videodisc system, proceduralized lessons, and lessons addressing STE/ICE.
Negative:	The device frequently malfunctioned and operated too slowly.
Experiment 5 (MOS 24C10) Seville/Burtek Device	
Positive:	Respondents hold favorable opinions of the device.
	"Hands-on" troubleshooting is a highly valued device feature.
Negative:	Low fidelity components of the 3-D module reduced device effectiveness.
Experiment 6 (MOSs 24E, G, R) Grumman Device	
Positive:	Respondents hold favorable opinions of the device.
	The device is safer than operational equipment.
	Automated features (request help, pre-lesson check, feedback, malfunction insertion) are valuable.
	The device allows students to practice tasks that cannot be practiced on operational equipment.
Negative:	The device frequently malfunctioned.
	Rebooting is a poor method for restarting a lesson.
	The instructor CRT provides little valuable information.
	Lessons are inflexible.

Benefits of the Grumman Approach

The Grumman approach can be characterized as one that incorporates recent advances in microelectronics and video storage. Respondents at APG indicated that they were favorably impressed by the videodisc approach used in the Grumman device. The still and motion video frames presented on the student CRT were useful at both locations. The sound track of the videodisc system was especially useful at APG since the soldiers who used the device at APG possessed poor reading skills. The touch panel simplified the students' tasks and added to their enthusiasm for using the training device. The component location "games" included as part of the curriculum at Fort Bliss also seemed to heighten student interest in the simulator.

A second, and perhaps more important, benefit of the Grumman approach, is the ability of the device to train students to perform tasks that cannot be practiced on operational equipment. This benefit was especially important at Fort Bliss since it is highly useful for radar mechanics to practice troubleshooting high voltage problems.

Liabilities of the Grumman Approach

The predominant liability of the Grumman approach is the very low reliability of the training device. This low reliability prevented instructors from understanding how to operate the device (the effectiveness of operator training sessions was seriously limited by device malfunctions), caused numerous delays in training sessions, and adversely affected student and instructor attitude towards working with the device. Although it is acknowledged that the reliability of a breadboard device cannot be expected to be as great as the reliability of a production model, the device must function well enough to demonstrate its capabilities. If the reliability of the device cannot be improved significantly without substantial effort, then the utility of the device is questionable.

Opinions solicited from individuals at both locations indicate that the device is inflexible. This inflexibility is manifested in several ways: 1) it would be costly and time consuming to change the material that is presented by the videodisc system, 2) student progress through each segment more closely resembles lockstep training than it does self-paced training, 3) the order in which students participate in training segments cannot be readily changed, and 4) the call instructor feature cannot be disabled by an instructor while a segment is in progress.

The manner in which malfunctions are inserted and removed from the device is awkward and time consuming. For example, at the start of the weekly check procedures, an instructor is required to spend a considerable amount of time ensuring that controls on the 3-D module are set to correct positions. As a student progresses through a lesson, these controls are set to a variety of new positions requiring the instructor to correctly reset the controls again after the lesson has been completed. This procedure of setting and resetting controls wastes a great deal of instructors' time.

Benefits of the Seville/Burtek Approach

As was the case with the Grumman device, the Seville/Burtek device allows students to practice tasks that they could not practice on operational equipment. Although this feature is especially important for training students to perform high voltage tasks on the radar transmitter at Fort Bliss, it was also found to be appropriate at APG.

Data gathered from both locations on both 3-D modules indicated that the device was quite flexible. Instructional materials presented to students by the slide projector or the student CRT can be updated easily. Students may skip steps in a lesson (if the instructor chooses) and may complete lessons in any sequence desired by the instructor. Further, the halt on student error feature can be disabled by an instructor during the course of a lesson.

The Seville/Burtek device was found to be fairly reliable. Although difficulties with the slide projector unit were experienced, these difficulties did not seriously delay training sessions.

Insertion of malfunctions is simple and efficient. At the beginning of the weekly check procedure, for example, instructors spend a minimal amount of time setting controls on the 3-D module out of tolerance. As students progress through a lesson, they must set these controls back to specific correct settings. Thus, it is the student rather than the instructor who must expend effort setting controls to their correct positions.

Liabilities of the Seville/Burtek Approach

At APG (especially the 63B30 MOS), it seemed that there were too many stimuli that required student attention. Further, procedures for entering decisions on the student station were somewhat complicated. A different approach for presenting information to students and accepting information from students appears warranted. That is, the number of stimuli to which students must attend should be decreased.

The durability of the Seville/Burtek device should be increased. If the device is to incorporate an effective remove/replace capability, the helical coils, bolts, etc. must be "hardened" to withstand the rigors of normal use.

CONCLUSIONS

The data presented in the previous section of the report provide answers to the three questions posed earlier:

- Do the curricula associated with the devices provide adequate training?

- How does the training effectiveness of the AMTESS devices compare to the training effectiveness of conventional training?

- Is the AMTESS concept of modular training simulators feasible?

In response to the first question, the data suggest that the devices do provide adequate training for both mechanical and electronic maintenance tasks. In all cases, students trained on the AMTESS devices passed Army School proficiency requirements. Responses by students, instructors, and course developers indicate that students are prepared to troubleshoot malfunctions in the field following training on the AMTESS devices.

Concerning training effectiveness, quantitative data clearly indicate that the devices as currently configured are inferior to conventionl training methods. In terms of the percentage of performance test steps passed, the devices were 87% to 98% as effective as conventional training. Performance of students trained on the devices was considerably less effective in terms of the time required to complete tasks and the number of data collector interventions. Conventionally trained students required 52% and 86% as much time and a mean of 78% as many data collector interventions to finish the tasks. Thus, students trained on the devices were able to perform their tasks, but they required more time and more supervision than conventionally trained students. Device training was least effective for the least experienced students.

Although these findings about the relative training effectiveness of the AMTESS devices are valid, care should be taken when generalizing from these results. It should be remembered that students who were trained on the AMTESS devices were tested on operational equipment without prior familiarization training on the operational equipment. Further, students were only tested one time. The relative training effectiveness of the AMTESS devices may be quite different from the results reported herein when simulator-trained students are provided familiarization training on operational equipment prior to testing, or when simulator-trained students are tested on multiple occasions. The quantitative training effectiveness data must also be evaluated in light of qualitative data concerning relative device effectiveness. Instructors, students, and others stated that the devices were valuable because they provided training that could not be provided on operational equipment (high voltage problems on the radar transmitter, oil pump failure on the diesel engine, etc.). Since the job performance of students who were provided this training could not be assessed, there was no way to quantify the training value of this feature of the devices. However, this important aspect of the devices should not be overlooked when assessing device training effectiveness.

Regarding the final question of the feasibility of the AMTESS concept, the data presented in this report demonstrate that the concept of modular, reconfigurable, training simulators is viable. Each AMTESS device successfully trained soldiers to perform maintenance tasks in both mechanical and electronic applications. Qualitative data indicate that device users are enthusiastic about the simulators. Although the durability and reliability of both devices was less than satisfactory, it is anticipated that these problems can be rectified as development efforts continue.

Simulators that are developed from the prototypes can be effectively used as part-trainers to support conventional training. In their current

configuration, however, the devices should not be substituted for conventional training unless dangerous or inaccessible equipment precludes the use of routine instruction. Conventional training may be particularly important for inexperienced soldiers and remove/replace tasks. Since the field evaluation could only provide gross comparisons between the simulators and conventional training, more definitive conclusions about the devices cannot be made. Currently planned laboratory research, however, will lead to specific conclusions about the design and use of generic simulator systems for different tasks, types of trainee, and training strategies.

The utility of future field evaluations of AMTESS devices (or other training devices) can be enhanced if the following recommendations are followed:

- The objective of the evaluation should be clearly defined and understood by all parties involved in the effort.

- A clear line of communication should be established between the individuals conducting the evaluation and the individuals who control resources essential to the success of the evaluation (personnel, equipment, facilities, etc.).

- Individuals who control essential resources must understand the importance of experimental rigor to the success of the effort.

REFERENCES

Campbell, S. C., Stonge, J. R., Cothran, P. C., Anaya, C. P., & Sicuranza, A. J. (1980). Army maintenance training and evaluation simulation system (AMTESS). Final Report, Grumman Report No. TS-AMTESS-R-80-40. Bethpage, NY: Grumman Aerospace Corporation Training Systems.

Criswell, E. L., Unger, K. W., Swezey, R. W., & Hays, R. T. (1983). Army maintenance training and evaluation simulation system (AMTESS) development and device features. SAI Report No. SAI-83-03-178. McLean, VA: Science Applications, Inc.

Criswell, E. L., Swezey, R. W., Allen, J. A., & Hays, R. T. (1984). Human factors analysis of two prototype Army maintenance training and evaluation simulation system (AMTESS) devices. SAI Report No. SAI-84-11-178. McLean, VA: Science Applications, Inc.

Dybas, R. T. (1981). Army maintenance training and evaluation simulation system (AMTESS). Proceedings of the Interservice/Industry Training Equipment Conference.

Dybas, R. T. (1983). Army maintenance training and evaluation simulation system (AMTESS). Paper presented at International Conference on Simulators. London, England: Institute of Electrical Engineers.

Evans, R. A., & Mirabella, A. (1982). Evaluation of the Army maintenance training and evaluation simulation system (AMTESS). Proceedings of the Interservice/Industry Training Equipment Conference.

Garlichs, E. A., Miller, E. J., Davis, E. (1983). Army maintenance training and evaluation simulation system (AMTESS): Trainer engineering design report, Volumes I, II, and III. Seville Technical Report No. TR 83-14. Pensacola, FL: Seville Research Corporation.

Woelfel, J. C., Duffy, P. J., Unger, K. W., Swezey, R. W., Hays, R. T., & Mirabella, A. (1984). A critical review of front-end analysis procedures employed in the Army maintenance training and evaluation simulation system (AMTESS). SAI Report No. SAI-84-01-178. McLean, VA: Science Applications, Inc.